AFRL-OSR-VA-TR-2013-0622

GLOBAL BEHAVIOR IN LARGE SCALE SYSTEMS

JOSE MOURA

CARNEGIE-MELLON UNIVERSITY

12/05/2013
Final Report

DISTRIBUTION A: Distribution approved for public release.

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RSL
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
02-12-2013

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
15-Jun-10 - 14-Jun-13

4. TITLE AND SUBTITLE
Global Behavior in Large Scale Systems

5a. CONTRACT NUMBER
FA9550-10-1-0291

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jose M F Moura

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Carnegie Mellon University
Office of Sponsored Projects
5000 Forbes Avenue
Pittsburgh, PA 15213-3815

8. PERFORMING ORGANIZATION REPORT NUMBER
1150088

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Chicago Regional Office
230 South Dearborn, Room 380
Chicago, IL 60605-1595

10. SPONSOR/MONITOR'S ACRONYM(S)
ONRRO (N62880)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution A - Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We study the emergence of global behavior in large scale networks. The underlying motivating application is epidemics like computer virus spreading, for example, in wide campus local networks. We consider multiple classes of viruses, each type bearing their own statistical characterization – exogenous contamination, contagious propagation, and healing. The network state (distribution of nodes infected by each class in the network) is a jump Markov process, not necessarily reversible, making it a challenge to obtain its invariant distribution. By suitable renormalization, in the limit of a large network (number of nodes), we describe the macroscopic or emergent behavior of the network by the solution of a set of deterministic nonlinear differential equations. These nonlinear differential equations are obtained by mean field analysis of the microscopic random dynamics. We study the qualitative behavior of the nonlinear differential equations describing the mean field dynamics.

15. SUBJECT TERMS
Epidemics, mean field, Markov random processes, virus spreading, asymptotic analysis, global behavior, emergent behavior, qu

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Jose M F Moura
					19b. TELEPHONE NUMBER *(include area code)* 4129832684

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Global Behavior in Large Scale Systems

José M. F. Moura

Dep. of Electrical and Computer Engineering

Carnegie Mellon University, Pittsburgh, PA, USA 15213

ph:(412)268-6341, fax: (412)268-3890 (moura@ece.cmu.edu)

Final Report

Grant: FA9550-10-1-0291

Program Manager: Dr. Tristan Nguyen

tristan.nguyen@us.af.mil

Air Force Office of Science and Research

875 Randolph Street, Suite 325 Room 3112, Arlington, VA 22203

December 3, 2013

Abstract

This research attained two main achievements: 1) It characterized the large deviations performance of distributed inference by cooperating agents, under a variety of sensing and communications noise and failure conditions; and 2) It derived the emergent behavior in large networks of agents, i.e., the network macroscopic behavior, from the microscopic random interactions among the agents.

1 Introduction

In this research we considered two main problems: 1) large deviation error performance in distributed inference; and 2) emergent behavior in large networks of distributed interacting agents. On the first item, we studied the large deviations performance of a particular class of distributed algorithms we refer to as consensus+innovations distributed detection over *noisy* networks, where agents at a time step k cooperate with their immediate neighbors (consensus) and assimilate their new observations (innovation). We showed that, under noisy communication, *all agents* can still achieve an exponential error rate, even when certain (or most) agents cannot detect the event of interest in isolation. The key to achieving this is the appropriate design of the time-varying weight sequence $\{\alpha_k = b_0/(a+k)\}$ by which agents weigh their neighbors' messages. We found a communication payoff threshold on the communication noise power, i.e., the critical noise power below which cooperation among neighbors improves detection performance and above which the noise in the communication among agents overwhelms the distributed detector performance. Our studied quantified several tradeoffs among network parameters and between the time (or number of measurements) needed for reliable distributed decision and the transmission power invested by the agents. Section 2 describes the problem and main accomplishments; see our publications [1, 2, 3, 4, 5, 6, 7, 8, 9] for further details. Part of this work underlies the doctoral thesis [10].

On the second item, we studied the emergence of global behavior in large scale networks. The underlying motivating application was epidemics like computer virus spreading, for example, in a wide campus local networks. We considered multiple classes of viruses, each type bearing their own statistical characterization - exogenous contamination, contagious propagation, and healing. The network state (distribution of nodes infected by each class in the network) is a jump Markov process, not necessarily reversible, making it a challenge to obtain its invariant distribution. By suitable renormalization, in the limit of a large network (number of nodes), we described the macroscopic or emergent behavior of the network by the solution of a set of deterministic nonlinear differential equations. These nonlinear differential equations were obtained by mean field analysis of the microscopic random dynamics. We established the qualitative behavior of the nonlinear differential equations describing the mean field dynamics. Section 3 describes the problem and main accomplishments; see our publications [11, 12, 13, 14, 15] for further details. A doctoral

thesis on this topic, [16], is close to being finished.

2 Large Deviation Performance in Distributed Inference

We describe briefly the distributed inference problem in the context of *distributed* simple binary hypothesis testing: N agents cooperate, through a sparse, connected communications graph $G = (V, E)$ (V is the set of agents, the nodes of the graph G, and E is the set of interagents channels, the edges of G) to decide at each time k, $k = 1, 2, \cdots$, between two possible states of nature, H_0 and H_1. This problem arises in many applications including classical surveillance, but now in a distributed setting like in netted, multisite, or MIMO radars, e.g., [17], where a system of spatially separated networked multistatic radar stations cooperate at every time $k = 1, 2, \cdots$, to detect the presence or absence of a target, or cognitive radio networks where distributed agents detect a primary user, e.g., [18]. Consider the following *distributed* sequential detector. At each time k, agent i executes three tasks: 1) makes an observation $y_i(k)$; 2) updates its local detection statistic $x_i(k)$ by a *distributed* algorithm:

$$x_i(k+1) = \underbrace{W^1_{ii}(k)x_i(k) + \sum_{j \in O_i} W^1_{ij}(k)x_j(k)}_{\text{consensus}} + \underbrace{W^2_i(k)\eta_i(k+1)}_{\text{innovations}} \tag{1}$$

where: $W^1_{ij}(k)$ and $W^2_i(k)$, $1 \leq i, j \leq N$, are weights; $\eta_i(k)$ is the local instantaneous log-likelihood ratio of agent i at time k computed from its own $y_i(k)$; and O_i is the set of neighbors of agent i as determined by the edge set E of the graph G; and 3) makes a decision by thresholding its detection statistic:

$$x_i(k) \ \overset{[}{H_0}]H_1 \gtrless \gamma, \ \ k = 1, 2, \cdots, \ i = 1, \cdots, N. \tag{2}$$

Equation (1) updates the test statistic with a two-step structure: the first, given by the first two terms on the right hand side (rhs) of (1), is like *consensus* and reflects the cooperation among agents – it averages the local statistic $x_i(k)$ of i with the local statistics $x_j(k)$ received from the neighbors $j \in O_i$; and the second, which we refer to as an *innovations* step, assimilates the measurement $y_i(k)$ through the instantaneous local log-likelihood $\eta_i(k)$. Hence, we refer to the local updating (1)

at each agent i as a *consensus+innovations* distributed algorithm and to the set of N detectors (1) and (2), $i = 1, \cdots, N$, as the consensus+innovations *distributed* detector, or distributed detector for short. We are fundamentally concerned with how 'good' can we make the distributed detector, i.e., what performance guarantees can we provide, when we carefully design the weight sequences $W_{ij}^1(k)$ and $W_i^2(k)$ in (1). To be more specific, we benchmark the error detection performance of the distributed detector with respect to the error performance of the Neyman-Pearson centralized sequential detector, which, under appropriate assumptions, is:

$$x(k) \overset{[}{H_0}]H_1 \gtrless \gamma, \quad k = 1, 2, \cdots \tag{3}$$

where the centralized log-likelihood ratio $x(k)$ is given by:

$$x(k) = \frac{1}{k} \sum_{i=1}^{N} \sum_{j=1}^{k} \eta_i(j), \tag{4}$$

and $\eta_i(k)$ is a renormalization of the local instantaneous log-likelihood ratio computed by agent i at time k from its instantaneous observation $y_i(k)$. Our goal is to determine the conditions under which and then show that, by carefully designing the weights $W_{ij}^1(k)$ and $W_i^2(k)$ in (1), we can similarly guarantee exponential rate decay at *every* agent i by a distributed detector (1) and (2), i.e., the error probability of the distributed detector at each and every agent i, decays *asymptotically* exponentially fast.

We consider these design and performance guarantee questions under a fairly general setting that takes into consideration limitations that may not necessarily arise in a centralized setting but are natural in many distributed applications. Because of limited power, not only are 1) the observations $y_i(k)$ of agent i noisy, affected by *sensing* noise, but also 2) the communications among agent i and its neighboring agents (when they cooperate) are noisy, impacted by *communications* noise. Our research considered general nonlinear noises. Here we explain our results by considering Gaussian sensing *and* communication noises. Our research established exponential error rate of decay for distributed detection. Note that, with noisy communications, the updating of the local statistic at agent i does not follow equation (1) but is more complex, see for details [5]. Extension to (non-Gaussian) quantized inter-agent communication and to Gaussian temporally correlated

sensing and communication noises are in [6].

Brief comment on the literature. We place our results in the context of the literature. There is a vast literature on decentralized and distributed inference. While we consider a distributed architecture, i.e., with no fusion center, references [19, 20, 21, 22, 23] consider decentralized parallel fusion architectures, where all agents communicate with a fusion center. References [24, 25, 26, 27, 28] have a distributed architecture (no fusion center) but are of the consensus type–each sensor makes a single observation and then the sensors fuse their local decisions by the consensus algorithm, or by belief propagation like in [24]. Reference [29] and the algorithm in Section IV in [30] are essentially of the consensus type, since they run consensus till convergence between each round of measurements. The algorithm in Section V in [30] assumes a complete architecture, or, if not, it uses a multihop protocol, so that each sensor has access to the observations of *all* the sensors at *each* and *every* time step. These references stand in contrast with the class of algorithms we consider: we use a *consensus+innovations* algorithm, i.e., a distributed algorithm (no fusion center) that interleaves *consensus* with *innovations* (processing of the observations) at the *same* time step, rather than running consensus to convergence in *between* successive observations.

We now contrast our work with [31, 32, 33, 18, 34, 35, 36, 37, 38, 39, 40, 41] that, like ours, are distributed, include communication among neighbors, and process the new observations at every time step as they are measured. We first comment that the main features that distinguish our work from these works are: 1) we consider *single scale* distributed detectors; 2) the communications among agents is corrupted by additive noise; and 3) we are primarily concerned with showing exponential error rate (with appropriate choice of the weights $W^1_{ij}(k)$, $W^2_i(k)$ in (1).) References [32, 33, 18] look at distributed LMS and RLS adaptive algorithms. They assume noiseless communications among agents (no additive noise) and they do not study the decay rate of the error probability[1]. Reference [35] addresses the problem of distributed change detection (a tracking type of problem) allowing for random averaging matrices and spatio-temporally correlated data, but this work does not consider noise in the communication among agents, nor is it concerned with establishing the exponential error rate of the algorithm therein. References [36, 2, 38] consider link failures but no additive noise in the intra-agents' communication. Also, [36] considers

[1]Coupling [18] with the results in [34, 31] that considers diffusion estimators with additive communication noise, the probability of error of the LMS detector in [18] does not go to zero as the number of observations grows to infinity, let alone achieve exponential decay rate, in contrast with the performance of our distributed detector.

the limiting behavior of their distributed detector when the difference between the means under the two hypotheses goes to zero, a very different problem from the problem we considered in our research. Our early work [39] considers deterministically time varying networks and no communications noise. Our work in [40] is concerned with estimation and considers a very general model that includes agent failures, link failures, and various degrees of either quantized or noisy communications. Because this reference studies estimation and not detection, it is not concerned with exponential decay rates of the error probability as we considered in our research in this project; rather, it shows consistency, asymptotic efficiency, and normality of the estimates through stochastic approximation and Lyapounov function arguments and through bounding pathwise behavior, rather than through large deviations arguments as we apply here to our detection analysis. A nonlinear estimator in [40] is *mixed* scale, while the class of detectors we study in this paper is *single* scale. The corresponding mixed scale algorithms for detection are presented and studied in [41], which, to the best of our knowledge, and within the consensus+innovations *detection* literature, is, like in our wrk on distributed detection, the only reference to consider additive noise in the communications among agents (also, with no link failures.) Our results contrast with [41], for the distributed sequential detector that we design, we establish that the error probability *at each agent* decays exponentially fast; we demonstrate this under broad conditions, including unequal local agents' sensing signal-to-noise ratios and when certain or most agents are locally not detectable.

In our references [2, 3], we focus on how link failures impinge on detection performance, while in [5] we show that *additive communication noise* in the links impacts in a qualitatively different way the error performance; with link failures, more communication among agents can only improve the error performance, since when communication does happen agents receive their neighbors detection statistics unencumbered by noise. But with additive communication noise, a clear tradeoff arises between communication noise and amount of information flow (or how often agents communicate;) this leads to a phase change behavior: only when the communication noise power is below a threshold does increased or more often cooperation improve performance–in that the distributed error performance of the worst (noisiest) agent is better than the isolated (no cooperation) performance of the best agent. While in [2, 3] we model certain averaging matrices as independent identically distributed (i.i.d.) so that their distribution is time invariant, in [5], because of time-decaying weights, the corresponding weight matrices are time varying, forcing us

to develop new analysis to show asymptotic stability of certain time varying systems. We refer to our publications [1, 2, 3, 4, 5, 6, 7, 8, 9] for additional details on the specific large deviation performance results and tradeoffs we obtained.

3 Emergent Behavior in Networks of Interacting Agents

Many complex dynamical systems exhibit *emergent behavior* – a well-structured macroscopic dynamics induced by simple, possibly random, local rules of interacting agents. Flocks of birds, ant colonies, beehives, brain neural networks, invasive tumor growth, and epidemics are all examples of large scale interacting agents systems displaying complex adaptive functional behaviors. Under appropriate initial conditions, a flock of birds reaches consensus on its cruise velocity while each bird probes only its nearest neighbors dynamics without a preferred leader in the flock (refer to [42]). This gives rise to synchronized flocking flying formations. Ant colonies can design optimal trails to access sources of food even though no ant bears the cognitive ability to shape up the colony to its blueprint mature optimal global behavior. Roughly, each scout-ant wanders around randomly tracking the leftover pheromone released by its scout peers. Reference [11] establishes the emergent dynamics of an idealized stochastic network model for ant colonies as the fluid limit dynamics of the network model (as the colony grows large). Seizure is an intricate outcome of the complex neural network dynamics of the brain. Reference [43] presents an overview of graphical dynamical models that have been applied to better understand the nature of seizures and bridge the microscopical electrical activity in the brain with the clinical observations of the phenomenon.

Our work has two main dimensions: 1) *Mean field dynamics.* The first dimension derives through analysis of the random interactions among the population agents (e.g., virus) a set of nonlinear differential equations that describes the emergent dynamics in the limit of large population sizes. 2) *Qualitative behavior.* The second dimension considers the qualitative behavior of these nonlinear differential equations to address questions like in the limit of large time are there particular strains of virus that survive, or what fraction of the population is infected by what virus. In both of these questions we are concerned with the impact of the network topology that captures the local interaction among population individuals.

Mean field dynamics. The challenge in studying large scale systems lies in their **high-**

dimensionality plus the **coupling** among the agents via their interactions. Together, these are the needed ingredients to induce emergent behavior. For instance, consider N agents whose state-vector

$$\mathbf{X}^N(t) := (X_1(t), \ldots, X_N(t))$$

evolves as a jump Markov process over the state space

$$\mathcal{S}^N := \{0,1\}^N .$$

If the agents are independent, then it turns out that the state of each agent evolves as a jump Markov process and, moreover, any state construct

$$(f\,(X_1(t), \ldots, X_N(t)))\,,$$

where $f : \{0,1\}^N \to \mathbb{R}^M$ bears appropriate measurability properties (we skip the details here), is a Markov jump process. For instance, the fraction of agents at state 1,

$$f\,(X_1(t), \ldots, X_N(t)) = \sum_{i=1}^{N} X_i(t)/N,$$

is Markov. Even for large N, due to the independence assumption, a qualitative analysis of $\left(X^N(t)\right)$ becomes tractable, but, in this example of independent agents, any weak law of large numbers will reflect the average behavior of each individual agent rather than an emergent global cooperative behavior. When the agents are coupled – e.g., an agent switches to state 1 with a rate that is proportional to the number of its neighbors in state 1–then, in general, neither the state of each agent is Markov nor the *macroscopic* low-dimensional states

$$(f\,(X_1(t), \ldots, X_N(t)))$$

are Markov and studying the *microscopic* high-dimensional dynamical system $\left(\mathbf{X}^N(t)\right)$ becomes quickly unfeasible with the number of agents N. Establishing the emergent dynamics or, in other

words, the functional weak law of large numbers under an arbitrary coupling topology of the agents is challenging. For the special case of a *complete* topology of interaction–any agent evenly affects any other agent in the cloud–low-dimensional *macroscopic* state-variables may still be Markov, even though the state of each individual agent is no longer Markov. Again, for complete networks, the fraction of infected nodes

$$f\left(X_1(t),\ldots,X_N(t)\right)=\sum_{i=1}^{N} X_i(t)/N$$

is Markov. Under this complete network setting, the emergent behavior is framed as the fluid limit dynamics of a global state variable

$$(\mathbf{Y}(t)) := \left(f\left(X_1(t),\ldots,X_N(t)\right)\right)$$

of interest. For example, reference [44] considers a multiclass flow of packets over a **complete** network with finite capacity nodes. It defines the macroscopic state variable

$$\left(\mathbf{Y}^N(t)\right)=\left(Y_1^N(t),\ldots,Y_L^N(t)\right)$$

that collects the fraction of nodes $Y_i^N(t)$ with a particular distribution i of packets over the different classes. Reference [44] proves that the empirical distribution

$$\left(\mathbf{Y}^N(t)\right)$$

converges weakly, with respect to the Skorokhod topology on the space of sample paths, to the solution of a vector ordinary differential equation.

For general topologies, the evolution of *macroscopic* state variables is intricately tied to the high-dimensional *microscopic* state $\left(\mathbf{X}^N(t)\right)$ of the system. Reference [45] proposes to consider the impact of the topology on the diffusion of a virus in the network, but, to overcome the coupling difficulty that arises with non complete networks, reference [45] departs from a peer-to-peer diffusion model. The authors in [45] replace the exact transition rates of the microstate process $(\mathbf{X}(t))$ by their average to establish their N-intertwined model. Were the states of the nodes independent

processes (a very strong assumption) and the resulting N-intertwined model would be an exact model to describe the dynamics of the likelihood of infection of each node as pointed out by the authors.

In our work, we went beyond the complete network model to establish the **exact** meanfield dynamics of a multi-virus epidemics over the class of multipartite networks, without making any artificial simplifying assumptions. We assume in our work a stochastic network model for the peer-to-peer spread of different strains of virus among a cloud of agents and establish the emergent dynamics of the epidemics. The emergent behavior is the fluid limit dynamics of the fraction of infected nodes over time. Namely, our work shows that, when the number of agents goes to infinity in a certain structured way, the fraction of infected agents at each *island* in the multipartite network converges weakly to the solution of a set of nonlinear ordinary differential equations.

This work established the *macroscopic scale* dynamics of a multi-virus epidemics or diffusion over large stochastic *non-complete* networks of agents.

Qualitative behavior. The second type of questions of interest that we addressed included when does a virus persists, when among multiple strains of virus we observe survival of the fittest, or what is the distribution of the fraction of infected agents over the various strains of virus in the network. These are well studied when the network is *complete*, i.e., any agent interacts directly with any other agent, and a vast body of literature describes the dynamics of the fraction of infected nodes by nonlinear ordinary differential equations (ODEs) that, as noted above, are arrived at through conservation or full mixing arguments, [46]. As also noted above, these nonlinear ODEs can also be rigorously derived when the network is complete because the fraction of infected nodes in the complete network is a Markov process under the standard independence assumptions on the peer-to-peer (*microscopic*) infection process, and the resulting *macroscopic* or *global behavior* of the epidemics is the fluid limit of this Markov process as the size of the complete network grows to infinity, see [47, 44]. When the network is not complete, the fraction of network infected nodes is no longer Markov and studying the network global or macroscopic behavior is the challenge we addressed in the previous paragraph. The mean field equations we obtained with our analysis are nonlinear coupled ODEs. We then studied the qualitative behavior of these mean field ODEs, i.e., the stability of their equilibria dynamics, to establish the emergent network macroscopic behaviors. Their coupled nonlinear behavior defies the use of Lyapunov methods. We developed a

new methodology that upper- and lower-bounds the limiting dynamics of the stochastic network by the much simpler to analyze dynamics of first order nonlinear systems. We considered single- and multi-virus epidemics and arbitrary regular multipartite networks.

Summary. Our work reported in [11, 12, 13, 14, 15] derives rigorously from basic peer-to-peer principles of diffusion the characterization of the global diffusion or infection behavior in multipartite networked systems in the limit of large systems. Our work is a *microscopic-to-macroscopic* study that goes beyond complete networked systems to obtain the **exact** impact of a non-complete topology on global infection and diffusion dynamics.

4 Main Conclusions

We now present briefly the main conclusions of our work. We consider separately the two main classes of results we obtained. **Large deviation performance results for distributed inference.** We designed a consensus+innovations distributed detector that achieves exponential error rate at all agents under noisy communication links, even when certain (or most agents) in isolation cannot perform successful detection. The key is the appropriate design of the consensus time-varying weights. We parameterized in terms of several network parameters a threshold on the communication noise power above which any agent that successfully detects the event in isolation still improves its performance through cooperation over noisy links, while below which not even the best agent can improve its detection performance by cooperation. We showed with numerical examples the significance on detection performance of tuning the weight sequence, showed communication payoff occurring already at a high noise level – and hence it is typically worthwhile to cooperate, and illustrate tradeoffs between the time to decision – time to reduce the error probability below a prescribed value – and the total transmission power. References [1, 2, 3, 4, 5, 6, 7, 8, 9] summarize our main results on large deviation performance analysis under broad conditions of sensing and communication noise and link failures.

Emergent behavior in random networks of interacting agents. There are three issues in determining the macroscopic behavior in stochastic networks: 1) Finding a Markovian macrostate, i.e., low dimensional functionals of the microstate $\mathbf{X}^N(t)$ that are Markov; 2) Deriving the equations for the dynamics of the macrostate in the limit of large networks–the mean field dynamics of

the macrostate; and 3) Studying the qualitative dynamics of the mean field.

The first and second items are dealt in greater detail in [14]; the third is in [15]. We consider the two first items now.

1. **Mean field dynamics.** We established the fluid limit dynamics of a multivirus epidemics with K virus over a *multipartite* network of M islands from a peer-to-peer stochastic network model of diffusion. We proved that the normalized macrostate

$$\left(\overline{\mathbf{Y}}_{ij}^{\mathbf{N}}(t)\right)$$

   collecting the fraction of j-infected nodes $\left(\overline{Y}_{ij}^{\mathbf{N}}(t)\right)$ per island $i \in \{1, \ldots, M\}$ with $j \in \{1, \ldots, K\}$ over the network given by the graph $G^{\mathbf{N}}$ converges weakly, under the Skorokhod topology on the space of *càdlàg* sample paths, to the solution $(y(t))$ of a $(M \times K)$-dimensional ordinary differential equation. To this effect, we first proved that the underlying martingale perturbation $\left(\overline{M}^{\mathbf{N}}(t)\right)$ vanishes as $\mathbf{N}$ grows large, which implies that the macrostate family $\left(\overline{\mathbf{Y}}_{ij}^{\mathbf{N}}(t)\right)$ is tight in $\mathbf{N}$. Then, we showed that any weak accumulation point of the family $\left(\overline{\mathbf{Y}}^{\mathbf{N}}(t)\right)$ is solution to a vector ordinary differential equation with Lipschitz vector field. From the uniqueness of the solutions of the resulting meanfield differential equation, we concluded that the whole sequence $\left(\overline{\mathbf{Y}}_{ij}^{\mathbf{N}}(t)\right)$ converges weakly to the solution of this meanfield differential equation.

   We now consider the third item.

2. **Qualitative behavior.** We analyzed the limiting (in the number of nodes) dynamics of a virus spreading in a regular multipartite network. Our method to derive the qualitative analysis of such coupled nonlinear dynamical system is not Lyapunov theory nor numerical simulations based. Instead, we explored a monotonous structure of the system, upper/lower bounding by simpler solutions any solution of the mean field equations. Our main conclusions for symmetric generic regular multipartite networks are:

   (a) *Virus Resilience*: If the inter island infection parameter $\gamma > \frac{1}{d}$, where d is the island degree, the virus persists in the network; otherwise, it dies out.

(b) *Natural Selection–Survival of the Fittest*: Only one strain (the most virulent one) survives, the remaining weaker ones die out; if there is a strain $k^\star$ such that its infection rate is such that $\gamma^{k^\star} > \gamma^k$ for all $k \neq k^\star$ with $\gamma^{k^\star} > \frac{1}{d}$, where d is the island degree, then virus $k^\star$ persists in the network and all the remaining strains die out.

For general multipartite networks, the break of symmetry may defy natural selection; this is bing pursued in future research.

References [11, 12, 13, 14, 15] detail our approach and results on the topic of emergent behavior in large networks of agents.

References

[1] D. Bajovic, D. Jakovetic, J. Xavier, B. Sinopoli, and J. Moura, "Distributed detection over time varying networks: Large deviations analysis," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, 2010, pp. 302–309.

[2] ——, "Asymptotic performance of distributed detection over random networks," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

[3] D. Bajovic, D. Jakovetic, J. Xavier, B. Sinopoli, and J. Moura, "Distributed detection via gaussian running consensus: Large deviations asymptotic analysis," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4381–4396, 2011.

[4] D. Bajovic, D. Jakovetic, J. Moura, J. Xavier, and B. Sinopoli, "Large deviations analysis of consensus;innovations detection in random networks," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, 2011, pp. 151–155.

[5] D. Jakovetic, J. M. F. Moura, and J. ao Xavier, "Distributed detection over noisy networks: Large deviations analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, August 2012.

[6] D. Bajovic, D. Jakovetic, J. Moura, J. Xavier, and B. Sinopoli, "Large deviations performance of consensus+innovations distributed detection with non-gaussian observations," *Signal Processing, IEEE Transactions on*, vol. 60, no. 11, pp. 5987–6002, 2012.

[7] D. Jakovetic, J. Moura, and J. Xavier, "Consensus+innovations detection: Phase transition under communication noise," in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, 2012, pp. 1559–1563.

[8] D. Bajovic, J. Xavier, J. Moura, and B. Sinopoli, "Exact rate for convergence in probability of averaging processes via generalized min-cut," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, 2012, pp. 2718–2725.

[9] ——, "Consensus and products of random stochastic matrices: Exact rate for convergence in probability," *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2557–2571, 2013.

[10] D. Jakovetic, "Distributed optimization: Algorithms and convergence rates," PhD, Carnegie Mellon University, Pittsburgh PA, USA, May 2013, department of Electrical and Computer Engineering.

[11] S. Kar and J. M. F. Moura, "Global emergent behaviors in clouds of agents," in *2011 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011, pp. 5796 –5799.

[12] A. Santos and J. M. F. Moura, "Emergent behavior in large scale networks," in *2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, December 2011, pp. 4485 –4490.

[13] ——, "Diffusion and topology: Large densely connected bipartite networks," in *2012 51th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, December 2012.

[14] A. Santos, J. M. F. Moura, and J. M. F. Xavier, "Emergent behavior in multipartite large networks: Multi-virus epidemics," 2013, to be submitted.

[15] ——, "Epidemics in multipartite networks: Emergent dynamics," July 2013, submitted.

[16] A. Santos, "Diffusion processes over complex networked systems," PhD, Carnegie Mellon University, Pittsburgh PA, USA, 2014, department of Electrical and Computer Engineering. Under progress.

[17] T. Derham, S. Doughty, K. Woodbridge, and C. Baker, "Design and evaluation of a low-cost multistatic netted radar system," *Radar, Sonar Navigation, IET*, vol. 1, no. 5, pp. 362 –368, October 2007.

[18] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 1917–1932, May 2011.

[19] J. N. Tsitsiklis, "Decentralized detection," *Adv. Statis. Signal Processing*, vol. 2, pp. 297–344, 1993.

[20] R. S. Blum, S. A. Kassam, and H. V. Poor, "Decentralized detection with multiple sensors: Part II–advanced topics," *Proc. IEEE*, vol. 85, pp. 64–79, January 1997.

[21] R. Viswanatan and P. R. Varshney, "Decentralized detection with multiple sensors: Part I–fundamentals," *Proc. IEEE*, vol. 85, pp. 54–63, January 1997.

[22] J. F. Chamberland and V. Veeravalli, "Decentralized dectection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407–416, February 2003.

[23] S. A. Aldosari and J. M. F. Moura, "Detection in sensor networks: the saddlepoint approximation," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 327–340, January 2007.

[24] V. Saligrama, M. Alanyali, and O. Savas, "Distributed detection in sensor networks with packet losses and finite capacity links," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4118–4132, November 2006.

[25] S. Kar and J. M. F. Moura, "Consensus based detection in sensor networks: topology design under practical constraints," in *Proc. IEEE Workshop Inf. Theory in Sensor Networks*, Santa Fe, NM, June 2007.

[26] S. Kar, S. A. Aldosari, and J. M. F. Moura, "Topology for distributed inference on graphs," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2609–2613, June 2008.

[27] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 355–369, January 2009.

[28] D. Jakovetic, J. Xavier, and J. M. F. Moura, "Weight optimization for consensus algorithms with correlated switching topology," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3788–3801, July 2010.

[29] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *CDC '07, 46th IEEE Conference on Decision and Control*, New Orleans, LA, Dec. 2007, pp. 5492–5498.

[30] R. Rahman, M. Alanyali, and V. Saligrama, "Distributed tracking in multihop sensor networks with communication delays," *IEEE Trans. Sig. Process.*, vol. 55, no. 9, pp. 4656–4668, September 2007.

[31] X. Zhao and A. H. Sayed, "Diffusion adaptation over networks under imperfect information exchange and non-stationary data," *to appear in IEEE Trans. Sig. Process.*, 2012, DOI: 10.1109/TSP.2012.2192928.

[32] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks based on diffusion estimation schemes," in *Proc. IEEE SPAWC '09, 10th IEEE International Workshop on Signal Processing Advances in Wireless Communications*, Perugia, Italy, June 2009, pp. 61–65.

[33] ——, "Diffusion LMS-based detection over adaptive networks," in *Proc. 39th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, October 2009, pp. 171–175.

[34] A. Khalili, M. A. Tinati, A. Rastegarnia, and J. A. Chambers, "Steady-state analysis of diffusion LMS adaptive networks with noisy links," *to appear in IEEE Transactions on Signal Processing*, 2012, DOI: 10.1109/TSP.2011.2173338.

[35] S. Stankovic, N. Ilic, M. S. Stankovic, and K. H. Johansson, "Distributed change detection based on a consensus algorithm," *IEEE Transactions o Signal Processing*, vol. 59, no. 12, pp. 5686–5697, December 2011.

[36] P. Braca, S. Marano, V. Matta, and P. Willet, "Asymptotic optimality of running consensus in testing binary hypothesis," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 814–825, February 2010.

[37] D. Bajovic, D. Jakovetic, J. Xavier, B. Sinopoli, and J. M. F. Moura, "Asymptotic performance of distributed detection over random networks," in *ICASSP 2011, IEEE International Conference on Acoustics, Speech and Signal Processing*, Prague, Chech Republic, May 2011, pp. 3008–3011.

[38] ——, "Distributed detection via Gaussian running consensus: large deviations asymptotic analysis," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4381–4936, September 2011.

[39] ——, "Distributed detection over time varying networks: large deviations analysis," in *44th Annual Allerton Conf. on Comm., Control, and Comp.*, Monticello, IL, September 2010, pp. 302–309.

[40] S. Kar, J. M. F. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication," *IEEE Transactions on Information Theory*, 2012, accepted for publication. [Online]. Available: arXiv:0809.0009v1 [cs.MA]

[41] S. Kar, R. Tandon, H. V. Poor, and S. Cui, “Distributed detection in noisy sensor networks,” in *Proc. IEEE ISIT 2011, International Symposium on Information Theory*, Saint Petersburgh, Russia, August 2011.

[42] F. Cucker and S. Smale, “Emergent behavior in flocks,” *IEEE Transactions on Automatic Control*, vol. 52, no. 5, pp. 852 –862, May 2007.

[43] M. P. Richardson, “Large scale brain models of epilepsy: dynamics meets connectomics,” *Journal of Neurology, Neurosurgery, and Psychiatry*, no. 83, pp. 1238–1248, December 2012.

[44] N. Antunes, C. Fricker, P. Robert, and D. Tibbi, “Stochastic networks with multiple stable points,” *The Annals of Applied Probability*, vol. 36, no. 1, pp. 255–278, 2008.

[45] P. Van Mieghem, J. Omic, and R. Kooij, “Virus spread in networks,” *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1 –14, Feb. 2009.

[46] D. J. Daley and J. Gani, *Epidemic Modelling: An Introduction*. Cambridge, UK: Cambridge University Press, 2001.

[47] N. Antunes, C. Fricker, P. Robert, and D. Tibbi, “Analisys of loss networks with routing,” *The Annals of Applied Probability*, vol. 16, no. 4, pp. 2007–2026, 2006.